

# A massively parallel nonoverlapping additive Schwarz method for discontinuous Galerkin discretization of elliptic problems

Maksymilian Dryja<sup>1</sup> · Piotr Krzyżanowski<sup>1</sup>

Received: 28 October 2014 / Revised: 6 February 2015 / Published online: 1 April 2015  
© The Author(s) 2015. This article is published with open access at Springerlink.com

**Abstract** A second order elliptic problem with discontinuous coefficient in 2-D or 3-D is considered. The problem is discretized by a symmetric weighted interior penalty discontinuous Galerkin finite element method with nonmatching simplicial elements and piecewise linear functions. The resulting discrete problem is solved by a two-level additive Schwarz method with a relatively coarse grid and with local solves restricted to subdomains which can be as small as single element. In this way the method has a potential for a very high level of fine grained parallelism. Condition number estimate depending on the relative sizes of the underlying grids is provided. The rate of convergence of the method is independent of the jumps of the coefficient if its variation is moderate inside coarse grid substructures or on local solvers' subdomain boundaries. Numerical experiments are reported which confirm theoretical results.

**Mathematics Subject Classification** 65N55

## 1 Introduction

In this paper we consider a second order elliptic equation

$$-\operatorname{div}(\rho \nabla u) = f,$$

---

✉ Piotr Krzyżanowski  
p.krzyzanowski@mimuw.edu.pl

Maksymilian Dryja  
m.dryja@mimuw.edu.pl

<sup>1</sup> University of Warsaw, Warsaw, Poland

with homogeneous Dirichlet boundary condition, where the diffusion coefficient  $\rho$  is a discontinuous function. The problem is discretized by a symmetric weighted interior penalty discontinuous Galerkin (DG) finite element method with simplicial nonmatching elements and piecewise linear functions. Our goal is to design and analyze a two-level nonoverlapping additive Schwarz method (ASM), see e.g. [16], for solving the resulting discrete problem with rate of convergence independent of the jumps of the coefficient and with a potential for a very high level of fine grained parallelism.

Usually, two level ASMs for discretizations on fine mesh of size  $h$  are defined by introducing a partitioning of the domain into subdomains of size  $H > h$ , where local solvers are applied in parallel. A global coarse problem is then based on the same partitioning. This method has recently been generalized for nonoverlapping domain decomposition methods for DG discretizations, allowing the coarse grid with mesh size  $\mathcal{H} \leq H$  to be a refinement of the original partitioning into subdomains where local solvers are applied—see e.g. [1, 14].

In the present paper, we take a different approach and consider local solvers' subdomains of size  $H$  and a coarse grid of size  $\mathcal{H}$  such that  $\mathcal{H} \geq H$ . Therefore, in an extreme case our local solver subdomain can be as small as a single element of the fine mesh (so then  $H = h$ ); on the other extreme, it can be set equal to the coarse grid cell,  $H = \mathcal{H}$ , the usual approach. By allowing small subdomains we substantially increase the level of parallelism of the method. Very small and cheap to solve local systems come in huge quantities, which can be an advantage on new multi-threaded processors. Moreover, small subdomains give more flexibility in assigning them to processors for load balancing in coarse grain parallel processing. In this way, an additional level of domain partitioning gives the user more parameters to fine tune the actual parallel performance, and thus overall efficiency, of the preconditioner for a given hardware architecture. The increased level of parallelism affects the condition number of the preconditioned system in a controlled way which we prove is, roughly speaking, of order  $O(\mathcal{H}^2/hH)$ . Numerical experiments reported in this paper confirm the bound is sharp in terms of mesh sizes.

In particular, if  $H = h$  we get the condition number of order  $O(\mathcal{H}^2/h^2)$  which then gradually improves, as  $H$  increases, to well-known  $O(\mathcal{H}/h)$  when  $H = \mathcal{H}$ , see [14]. It is also not surprising that if all partitionings scale in the same way so that  $\mathcal{H}/H$  and  $H/h$  are kept constant, then the condition number remains bounded independently of the actual values of  $\mathcal{H}$ ,  $H$ ,  $h$ .

The method discussed in the paper makes the preconditioned system condition number independent of the jumps of diffusion coefficient  $\rho$ , under the assumption that the distribution of the coefficient satisfies certain assumptions. We identify two such cases: when the variation of the coefficient is mild inside cells of the coarse grid, or when the coefficient is close to a constant on the skeleton of the partitioning into local solvers' subdomains.

The ASM discussed here is a generalization of [11] and, for continuous coefficient, of [1, 14]. See also [3] for a similar concept for continuous FE discretization where functions of the coarse space were constant inside substructures on which the original region was partitioned. The authors are unaware of other results concerning the influence of the sizes of local solver subdomains and the coarse grid on the condition

number when the former grid is a refinement of the latter, in the context of DG methods for the discretization of elliptic problems with discontinuous coefficient. Other recent works towards domain decomposition preconditioning of DG discretizations of problems with strongly varying coefficients include [2, 5, 6].

The paper is organized as follows. In Sect. 2, differential and discrete DG problems are formulated. In Sect. 3, ASM for solving the discrete problem is designed and analyzed. Numerical experiments are presented in Sect. 4.

For nonnegative scalars  $x, y$ , we shall write  $x \lesssim y$  if there exists a positive constant  $C$ , independent of  $x, y$  and mesh parameters  $h, H, \mathcal{H}$ , and of jumps of the diffusion coefficient  $\rho$  as well, such that  $x \leq Cy$ . If both  $x \lesssim y$  and  $y \lesssim x$ , we shall write  $x \simeq y$ .

## 2 Differential and discrete DG problems

Let  $\Omega$  be a bounded open polyhedral domain in  $R^d, d \in \{2, 3\}$ , with Lipschitz boundary  $\partial\Omega$ . We consider the following variational problem for given  $f \in L^2(\Omega)$  and  $\rho \in L^\infty(\Omega)$ :

Find  $u^* \in H_0^1(\Omega)$  such that

$$a(u^*, v) = (f, v)_\Omega, \quad \forall v \in H_0^1(\Omega), \quad (1)$$

where

$$a(u, v) = \int_\Omega \rho \nabla u \cdot \nabla v \, dx, \quad (f, v)_\Omega = \int_\Omega f v \, dx.$$

We assume that there exist constants  $\alpha_0$  and  $\alpha_1$  such that  $0 < \alpha_0 \leq \rho \leq \alpha_1$  a.e. in  $\Omega$  so that (1) is well-posed. Without loss of generality we shall additionally suppose that  $\alpha_0 \geq 1$ , which can always be guaranteed by simple scaling. This assumption will simplify the proof in Sect. 3.

In what follows we will analyze a preconditioner for a system of algebraic equations arising from a discretization of (1) with DG finite element method. The corresponding finite element spaces and the discrete problem are introduced below in the following subsection.

### 2.1 Finite element spaces and DG discretization

Let  $\mathcal{T}_h$  denote an affine, shape-regular but not necessarily conforming partition of  $\Omega$  by triangles in 2-D or tetrahedrons in 3-D,  $\mathcal{T}_h = \{K_1, \dots, K_{N_h}\}$ . We shall refer to  $\mathcal{T}_h$  as the “fine mesh”. The diameter of element  $K \in \mathcal{T}_h$  will be denoted by  $h_K$  (assumed uniformly bounded from above by an absolute constant) and we collect all elements’ diameters in a multi-parameter  $h = (h_{K_1}, \dots, h_{K_{N_h}})$ .

By  $\mathcal{E}_h^0$  we denote the set of all common (internal) faces (edges in 2-D) of elements in  $\mathcal{T}_h$ , so that  $e \in \mathcal{E}_h^0$  iff  $e = \partial K_i \cap \partial K_j$  is of positive measure. We will use symbol  $\mathcal{E}_h$  to denote the set of all faces (edges in 2-D) of fine mesh  $\mathcal{T}_h$ , that is those either in  $\mathcal{E}_h^0$  or on the boundary  $\partial\Omega$ . For  $e \in \mathcal{E}_h$  we set  $h_e = \text{diam}(e)$ .

We assume that  $\mathcal{T}_h$  is shape- and contact-regular in the sense of [9, Definition 1.38], that is, it admits a matching submesh  $\mathcal{T}_{\tilde{h}}$  which is shape-regular and such that for any  $K \in \mathcal{T}_h$  the ratios of  $h_K$  to diameters of simplices in  $\mathcal{T}_{\tilde{h}}$  covering  $K$  are uniformly bounded by an absolute constant. In consequence, if  $e = \partial K_i \cap \partial K_j \in \mathcal{E}_h^0$ , then

$$h_e \simeq h_{K_i} \simeq h_{K_j}. \quad (2)$$

Moreover, the number of neighboring elements in  $\mathcal{T}_h$  is bounded by an absolute constant.

For  $p \in \{0, 1\}$ , we denote by  $\mathcal{P}_p$  the set of polynomials of degree not greater than  $p$ . Then we define the finite element space  $V_h$  in which we will approximate (1),

$$V_h = \{v \in L^2(\Omega) : v|_K \in \mathcal{P}_1 \text{ for all } K \in \mathcal{T}_h\}. \quad (3)$$

Note that traces of functions from  $V_h$  are multi-valued on the skeleton  $\mathcal{E}_h$ .

We discretize (1) by the symmetric weighted interior penalty DG method, see for example [10, 13]:

Find  $u_h^* \in V_h$  such that

$$\mathcal{A}_h(u_h^*, v_h) = (f, v_h)_\Omega, \quad \forall v_h \in V_h, \quad (4)$$

where

$$\begin{aligned} \mathcal{A}_h(u, v) \equiv & \sum_{K \in \mathcal{T}_h} (\rho \nabla u, \nabla v)_K + \sum_{e \in \mathcal{E}_h} \langle \gamma[u], [v] \rangle_e \\ & - \sum_{e \in \mathcal{E}_h} (\langle [u], \{\rho \nabla v\}_\omega \rangle_e + \langle \{\rho \nabla u\}_\omega, [v] \rangle_e). \end{aligned}$$

Here for  $K \in \mathcal{T}_h$  and  $e \in \mathcal{E}_h$  we use standard notation

$$(u, v)_K = \int_K u v \, dx \quad \text{and} \quad \langle u, v \rangle_e = \int_e u v \, d\sigma.$$

In what follows we shall assume that  $\rho$  is piecewise constant on  $\mathcal{T}_h$  (equal to its average value on each element), thus  $\rho_i := \rho|_{K_i}$  is constant for  $i = 1, \dots, N_h$ . On  $e = \partial K_i \cap \partial K_j$  we set

$$\gamma = \frac{\delta}{h_e} \frac{\rho_i \rho_j}{\rho_i + \rho_j}, \quad \{\rho \nabla u\}_\omega = \omega_j \rho_i \nabla u_i + \omega_i \rho_j \nabla u_j, \quad [u] = u_i n_i + u_j n_j,$$

with  $\omega_j = \rho_j / (\rho_i + \rho_j)$ . The unit normal vector pointing outward  $K_i$  is denoted by  $n_i$ . On  $e$  which lies on the boundary of  $\Omega$  and belongs to a face of  $K_i$ , we set  $\{\rho \nabla u\}_\omega = \rho_i \nabla u_i$ ,  $[u] = u_i n_i$  and  $\gamma = \delta \rho_i / h_e$ .

For sufficiently large  $\delta$  the discrete problem (4) is well-defined, according to the following lemma:

**Lemma 1** ([10, 13]) *There exists positive  $\delta_0$  such that if  $\delta \geq \delta_0$  the bilinear form  $\mathcal{A}_h(\cdot, \cdot)$  is symmetric, positive definite, and spectrally equivalent to a simplified form*

$$A_h(u, v) = \sum_{K \in \mathcal{T}_h} (\rho \nabla u, \nabla v)_K + \sum_{e \in \mathcal{E}_h} \langle \gamma[u], [v] \rangle_e,$$

that is, there holds

$$\mathcal{A}_h(u, u) \simeq A_h(u, u) \quad \forall u \in V_h. \quad (5)$$

In what follows we shall take as granted that  $\delta$  is a fixed parameter such that  $\delta \geq \delta_0$ . The seminorm of a function  $f$  from the Sobolev space  $H^s(U)$  will be denoted by  $|f|_{H^s(U)}$ . For short, the  $L^2$ -norm of  $f$  will then be denoted by  $|f|_{0,U}$  or simply  $|f|_U$ . We also define the broken norm  $\|u\|_{1,h}$  by the identity  $\|u\|_{1,h}^2 = A_h(u, u)$ . Then there holds the following approximation result:

**Lemma 2** ([10, 13]) *Let  $u^* \in H_0^1(\Omega)$  be the solution of (1), and let  $u_h^* \in V_h$  satisfy the discrete problem (4). If  $u_{|K}^* \in H^2(K)$  for all  $K \in \mathcal{T}_h$ , then*

$$\|u^* - u_h^*\|_{1,h}^2 \lesssim \sum_{K \in \mathcal{T}_h} h_K^2 \rho_{|K} |u^*|_{H^2(K)}^2.$$

### 3 Additive Schwarz method with small subdomains

The condition number of the discrete problem (4) can be prohibitively large, affected by both the fine mesh size and by the magnitude of jumps in  $\rho$ . Thus, for an iterative solution of (4), some preconditioning is necessary. In this section we consider an ASM whose ingredients are: a large number of local solvers on relatively small nonoverlapping subdomains, and a coarse grid solver defined on a possibly much coarser grid.

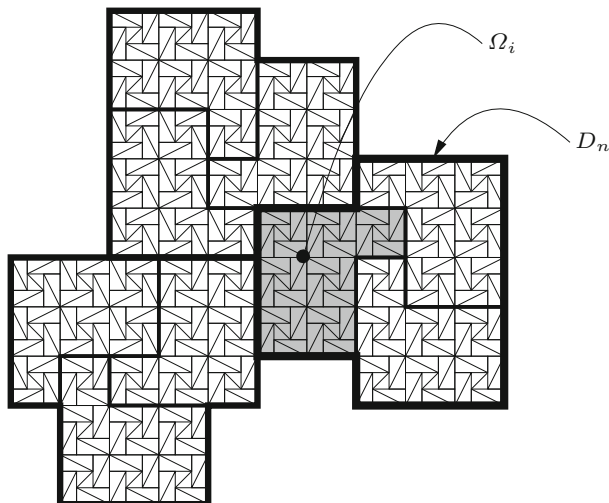
Let us therefore introduce  $\mathcal{T}_H$  as a partition of  $\Omega$  into  $N_H$  disjoint open polyhedral subdomains  $\Omega_i$ ,  $i = 1, \dots, N_H$ , such that  $\bar{\Omega} = \bigcup_{i=1, \dots, N_H} \bar{\Omega}_i$  and that each  $\Omega_i$  is a union of certain elements from the fine mesh  $\mathcal{T}_h$ , which we will denote  $\mathcal{T}_h(\Omega_i) = \{K \in \mathcal{T}_h : K \subset \Omega_i\}$ . We set  $H_i = \text{diam}(\Omega_i)$  and  $H = (H_1, \dots, H_{N_H})$ . We shall refer to this partition as the “local solvers’ grid”, because our ASM is going to solve (in parallel) local problems defined on the fine grid restricted to each of these subdomains.

Next, let  $\mathcal{T}_{\mathcal{H}}$  be a division of  $\Omega$  into  $N_{\mathcal{H}}$  disjoint open polyhedral regions  $D_n$ ,  $n = 1, \dots, N_{\mathcal{H}}$ , such that  $\bar{\Omega} = \bigcup_{n=1, \dots, N_{\mathcal{H}}} \bar{D}_n$  and that each  $D_n$  is a union of certain elements from the local solvers’ grid  $\mathcal{T}_H$ . We set  $\mathcal{H}_n = \text{diam}(D_n)$  and then  $\mathcal{H} = (\mathcal{H}_1, \dots, \mathcal{H}_{N_{\mathcal{H}}})$ . We shall call this partition the “coarse grid”.

We clearly have  $N_{\mathcal{H}} \leq N_H \leq N_h$  and

$$\mathcal{T}_{\mathcal{H}} \subseteq \mathcal{T}_H \subseteq \mathcal{T}_h$$

(inclusions understood in the sense of subsequent refinements of the coarsest partitioning), and  $\max h \leq \max H \leq \max \mathcal{H}$ . Note that in [1, 14] the inclusions come in a



**Fig. 1** Relationship between  $\mathcal{T}_h$ ,  $\mathcal{T}_H$ ,  $\mathcal{T}_{\mathcal{H}}$ . Thin lines correspond to the fine grid  $\mathcal{T}_h$ . Local solvers' grid  $\mathcal{T}_H$  is marked with lines of medium width; an example local solvers' subdomain  $\Omega_i$  is grayed. Coarse partition  $\mathcal{T}_{\mathcal{H}}$ —here consisting of three local solvers' subdomains per its element—is marked with thickest lines

different order:  $\mathcal{T}_H \subseteq \mathcal{T}_{\mathcal{H}} \subseteq \mathcal{T}_h$ , making the use of very small subdomains infeasible due to excessive increase of the coarse space dimension.

An example partitioning of  $\Omega$  into fine, coarse and local solvers' grids is shown in Fig. 1.

For  $S$  belonging either to  $\mathcal{T}_H$  or to  $\mathcal{T}_{\mathcal{H}}$  let us introduce the notation referring to the skeleton of the fine mesh restricted only to  $S$ :

$$\mathcal{E}_h(S) = \{e \in \mathcal{E}_h : e \subset \bar{S}\}$$

and its boundary and interior parts as well:

$$\mathcal{E}_h^\partial(S) = \{e \in \mathcal{E}_h(S) : e \subset \partial S\}, \quad \mathcal{E}_h^0(S) = \mathcal{E}_h(S) \setminus \mathcal{E}_h^\partial(S).$$

Finally, let us collect all fine grid faces belonging to the boundaries of elements of  $\mathcal{T}_H$  and  $\mathcal{T}_{\mathcal{H}}$ , respectively, to denote the skeletons of  $\mathcal{T}_H$  and  $\mathcal{T}_{\mathcal{H}}$ :

$$\mathcal{E}_H = \bigcup_{i=1}^{N_H} \mathcal{E}_h^\partial(\Omega_i) \quad \text{and} \quad \mathcal{E}_{\mathcal{H}} = \bigcup_{n=1}^{N_{\mathcal{H}}} \mathcal{E}_h^\partial(D_n).$$

We start the definition of the ASM by introducing a decomposition of  $V_h$ :

$$V_h = V_0 + \sum_{i=1}^{N_H} V_i \tag{6}$$

where the coarse space is

$$V_0 = \{v \in V_h : v|_{D_n} \in \mathcal{P}_0 \text{ for all } n = 1, \dots, N_{\mathcal{H}}\}$$

and for  $i = 1, \dots, N_H$  the local spaces are

$$V_i = \{v \in V_h : v|_{\Omega_j} = 0 \text{ for all } j \neq i\}. \quad (7)$$

Note that  $V_h$  is a direct sum of the local spaces. By choosing the lowest order coarse space, we reduce its dimension and, in consequence, the cost of solving the coarse problem.

Since now on, for  $\varphi$  defined on  $\Omega$ , if necessary we shall write  $\varphi_i$  to denote the restriction of  $\varphi$  to  $\Omega_i$ :

$$\varphi_i := \varphi|_{\Omega_i}.$$

Observe that this brings a change from the notation of Sect. 2.1.

Using decomposition (6) we define local operators  $T_i : V_h \rightarrow V_i$ ,  $i = 1, \dots, N_H$ , by “inexact” solvers

$$A_h(T_i u, v) = \mathcal{A}_h(u, v) \quad \forall v \in V_i,$$

so that on each subdomain one has to solve only a relatively small system of linear equations for  $u_i = T_i u|_{\Omega_i}$  such that for all  $v_i \in V_i$

$$\sum_{K \in \mathcal{T}_h(\Omega_i)} (\rho \nabla u_i, \nabla v_i)_K + \sum_{e \in \mathcal{E}_h^0(\Omega_i)} \langle \gamma[u_i], [v_i] \rangle_e + \sum_{e \in \mathcal{E}_h^\partial(\Omega_i)} \int_e \gamma u_i v_i = \mathcal{A}_h(u, v_i).$$

(For  $j \neq i$  we set  $(T_i u)|_{\Omega_j} = 0$ .) The coarse solve operator is  $T_0 : V_h \rightarrow V_0$  defined analogously as

$$A_h(T_0 u, v_0) = \mathcal{A}_h(u, v_0) \quad \forall v_0 \in V_0.$$

Note that on  $V_0$ , the approximate form  $A_h(\cdot, \cdot)$  coincides with  $\mathcal{A}_h(\cdot, \cdot)$  and simplifies to

$$A_h(u_0, v_0) = \sum_{e \in \mathcal{E}_{\mathcal{H}}} \langle \gamma[u_0], [v_0] \rangle_e \quad \forall u_0, v_0 \in V_0.$$

Finally, the preconditioned operator is

$$T = T_0 + \sum_{i=1}^{N_H} T_i. \quad (8)$$

In order to formulate the condition number estimate result for  $T$ , we make two additional assumptions, that the elements of partition  $\mathcal{T}_H$  do not differ too much in shape, as well as those in  $\mathcal{T}_h$ :

- (A1) There exists a reference simply-connected polyhedron  $\hat{D} \subset \mathbb{R}^d$  to which all coarse grid partition elements  $D_n \in \mathcal{T}_H$  are affinely homeomorphic and the aspect ratios of  $D_n$ ,  $n = 1, \dots, N_H$ , are uniformly bounded with constant independent of  $\mathcal{H}$ ,  $H$  and  $h$ . (An example of  $\mathcal{T}_H$  which satisfies this condition is a partition which forms a simplicial, shape-regular, not necessarily matching, partition of  $\Omega$ .)
- (A2) There exists a reference simply-connected, polyhedral structure  $\hat{\Omega} \subset \mathbb{R}^d$  to which all local solvers' grid subdomains  $\Omega_i \in \mathcal{T}_h$  are affinely homeomorphic and the aspect ratios of  $\Omega_i$ ,  $i = 1, \dots, N_H$ , are uniformly bounded with constant independent of  $\mathcal{H}$ ,  $H$  and  $h$ . Moreover, the number of neighboring regions in  $\mathcal{T}_h$  is uniformly bounded by an absolute constant  $\mathcal{N}$ . (An example of  $\mathcal{T}_h$  which satisfies this condition is a partition which forms a simplicial, shape-regular, not necessarily matching, locally uniform partition of  $\Omega$ .)

Assumptions (A1) and (A2) are required for Poincaré's and trace inequalities in the proof of Theorem 1.

**Theorem 1** *In addition to our general assumptions of Sect. 2, suppose that both (A1) and (A2) hold. Then  $T$  defined in (8) is symmetric with respect to  $\mathcal{A}_h(\cdot, \cdot)$  and*

$$\beta^{-1} \mathcal{A}_h(u, u) \lesssim \mathcal{A}_h(Tu, u) \lesssim \mathcal{A}_h(u, u) \quad \forall u \in V_h,$$

where

$$\beta = \max_{n=1, \dots, N_H} \frac{\mathcal{H}_n^2}{h H_n} \frac{R_n}{\underline{R}_n} + \max_{i=1, \dots, N_H} \frac{H_i}{\underline{h}_i} \frac{r_i}{\underline{r}_i}$$

with mesh-dependent parameters defined as

$$\begin{aligned} \underline{h}_i &= \min\{h_K : K \in \mathcal{T}_h(\Omega_i)\}, \\ \underline{h} H_n &= \min\{\underline{h}_i H_i : \Omega_i \subset D_n, i = 1, \dots, N_H\}, \end{aligned}$$

and coefficient-dependent parameters defined as

$$\begin{aligned} \underline{r}_i &= \min \rho_{|\Omega_i|}, & r_i &= \max\{\rho_{|e|} : e \in \mathcal{E}_h^\partial(\Omega_i)\}, \\ \underline{R}_n &= \min \rho_{|D_n|}, & R_n &= \max\{r_i : \Omega_i \subset D_n, i = 1, \dots, N_H\}. \end{aligned}$$

Therefore, the condition number of the preconditioned operator  $T$  is  $O(\beta)$ .

Before proceeding with the proof of Theorem 1, we state a useful trace inequality result for functions in  $V_h$  (see also [14], where the estimate of Lemma 3 is proved for a star-like region  $\Omega_i$ ; here we do not need this assumption):



**Lemma 3** *Under the assumptions of Sect. 2 together with (A2), for any  $i = 1, \dots, N_H$  there holds*

$$|v|_{0,\partial\Omega_i}^2 \lesssim \frac{1}{H_i} |v|_{0,\Omega_i}^2 + H_i \left( \sum_{K \in \mathcal{T}_h(\Omega_i)} |\nabla v|_{0,K}^2 + \sum_{e \in \mathcal{E}_h^0(\Omega_i)} \frac{1}{h_e} |[v]|_{0,e}^2 \right) \quad \forall v \in V_h.$$

*Proof* Let us fix  $i = 1, \dots, N_H$  and let us denote by  $\mathcal{T}_{\hat{h}}$  the fine mesh  $\mathcal{T}_h(\Omega_i)$  affinely mapped onto  $\hat{\Omega}$ . It suffices to prove that for  $v \in V_{\hat{h}}$ ,

$$|v|_{0,\partial\hat{\Omega}}^2 \lesssim |v|_{0,\hat{\Omega}}^2 + \|v\|_{1,\hat{h},\hat{\Omega}}^2 \quad (9)$$

on the reference domain  $\hat{\Omega}$ , where for short we denote

$$\|v\|_{1,\hat{h},\hat{\Omega}}^2 = \sum_{K \in \mathcal{T}_{\hat{h}}} |\nabla v|_{0,K}^2 + \sum_{e \in \mathcal{E}_{\hat{h}}^0(\hat{\Omega})} \frac{1}{\hat{h}_e} |[v]|_{0,e}^2,$$

and then apply the standard scaling argument. By (A2) and the assumptions on  $\mathcal{T}_h$  of Sect. 2.1,  $\mathcal{T}_{\hat{h}}$  is also shape- and contact-regular, with relevant constants independent of  $h, H, \mathcal{H}$ . Thus,  $\mathcal{T}_{\hat{h}}$  admits a shape-regular matching submesh  $\mathcal{T}_{\tilde{h}}$ , so it is sufficient to prove (9) on  $\mathcal{T}_{\tilde{h}}$ .

Our proof will use technical tools and results from [4]. Let us consider two subspaces of  $V_{\tilde{h}}$ : the space  $V_{\tilde{h}}^1$  of nonconforming  $P_1$  finite elements [7] on  $\mathcal{T}_{\tilde{h}}$ , and  $\tilde{W}_{\tilde{h}}$ —the continuous Lagrange finite element space consisting of piecewise polynomials from  $P_2$  (in 2-D) or  $P_3$  (in 3-D).

Let  $\mathcal{I} : V_{\tilde{h}} \rightarrow V_{\tilde{h}}^1$  and  $E : V_{\tilde{h}}^1 \rightarrow \tilde{W}_{\tilde{h}}$  be operators defined in [4, Eq. (2.1) and (3.1)]. Let us recall that  $\mathcal{I}v$  is defined by its values at the centers of the faces

$$(\mathcal{I}v)(c) = \frac{1}{|e|} \int_e \{v\} \, ds$$

where  $c$  denotes the center of face  $e$ . The enriching operator  $E$  is defined by the nodal values of  $v$ , so that in node  $p$ ,

$$(Ev)(p) = \frac{1}{n_p} \sum_{K \in \mathcal{V}(p)} v|_K(p),$$

where  $\mathcal{V}(p)$  is the set of elements in  $\mathcal{T}_{\tilde{h}}$  sharing node  $p$ , and  $n_p$  is the number of these elements.

We will use the following estimates from [4, Eq. (2.8), (2.9), (3.10)–(3.11) and estimates from Example 4.2]:

$$\sum_{K \in \mathcal{T}_{\tilde{h}}} |\nabla \mathcal{I}v|_{0,K}^2 \lesssim \|v\|_{1,\tilde{h},\hat{\Omega}}^2, \quad \sum_{K \in \mathcal{T}_{\tilde{h}}} \tilde{h}_K^{-1} |v - \mathcal{I}v|_{0,K}^2 \lesssim \|v\|_{1,\tilde{h},\hat{\Omega}}^2 \quad (10)$$

for all  $v \in V_{\tilde{h}}^1$ , and

$$|Ev|_{\hat{\Omega}}^2 \lesssim |v|_{\hat{\Omega}}^2, \quad |\nabla Ev|_{\hat{\Omega}}^2 \lesssim \sum_{K \in \mathcal{T}_{\tilde{h}}} |\nabla v|_{0,K}^2, \quad |Ev - v|_{\partial\hat{\Omega}}^2 \lesssim \sum_{K \in \mathcal{T}_{\tilde{h}}} \tilde{h}_K |\nabla v|_{0,K}^2 \quad (11)$$

for all  $v \in \tilde{W}_{\tilde{h}}$ .

For  $v \in V_{\tilde{h}}$  we have

$$|v|_{0,\partial\hat{\Omega}}^2 \lesssim |v - \mathcal{I}v|_{0,\partial\hat{\Omega}}^2 + |\mathcal{I}v - E(\mathcal{I}v)|_{0,\partial\hat{\Omega}}^2 + |E(\mathcal{I}v)|_{0,\partial\hat{\Omega}}^2. \quad (12)$$

By trace inequality and the last inequality of (10),

$$|v - \mathcal{I}v|_{0,\partial\hat{\Omega}}^2 \lesssim \sum_{K \in \mathcal{T}_{\tilde{h}}} \tilde{h}_K^{-1} |v - \mathcal{I}v|_{0,K}^2 \lesssim \|v\|_{1,\tilde{h},\hat{\Omega}}^2.$$

Using the last estimate of (11) and then the first from (10) we get

$$|\mathcal{I}v - E(\mathcal{I}v)|_{0,\partial\hat{\Omega}}^2 \lesssim \sum_{K \in \mathcal{T}_{\tilde{h}}} |\nabla \mathcal{I}v|_{0,K}^2 \lesssim \|v\|_{1,\tilde{h},\hat{\Omega}}^2.$$

Finally, as  $E(\mathcal{I}v) \in H^1(\hat{\Omega})$ , we can apply standard trace inequality and use (11) and (10), to conclude that

$$|E(\mathcal{I}v)|_{0,\partial\hat{\Omega}}^2 \lesssim |E(\mathcal{I}v)|_{0,\hat{\Omega}}^2 + |\nabla E(\mathcal{I}v)|_{0,\hat{\Omega}}^2 \lesssim |v|_{0,\hat{\Omega}}^2 + \|v\|_{1,\tilde{h},\hat{\Omega}}^2.$$

Substituting the above inequalities into (12) we arrive at (9).  $\square$

*Proof (of Theorem 1)* We follow the abstract theory of additive Schwarz methods, cf. [16, Theorem 2.7], and prove three key properties for  $T_0, T_1, \dots, T_{N_H}$ :

*Strengthened Cauchy–Schwarz inequality* It is straightforward to verify that if  $u \in V_i$  and  $v \in V_j$  with  $i, j \in \{1, \dots, N_H\}$  then  $\mathcal{A}_h(u, v) = 0$ , if only  $\Omega_i$  and  $\Omega_j$  are not neighbors, i.e. they do not share a face in  $\mathcal{E}_H^0$ . Thus, we conclude that then

$$\mathcal{A}_h(u, v) \leq \epsilon_{ij} \mathcal{A}_h(u, u)^{1/2} \mathcal{A}_h(v, v)^{1/2}$$

and the spectral radius of  $\epsilon = (\epsilon_{ij})_{i,j=1}^{N_H}$  is bounded by  $\mathcal{N}$ —the maximum number of neighbors of any element in  $\mathcal{T}_H$ . According to assumption (A2)  $\mathcal{N}$  is bounded independently of the mesh parameters  $h$  and  $H$  (and, clearly, independently of other parameters of the problem).

*Local stability* For all  $i = 0, 1, \dots, N_H$

$$\mathcal{A}_h(u, u) \leq \omega A_h(u, u) \quad \forall u \in V_i,$$

with absolute constant  $\omega$  independent of  $\rho, h, H$  and  $\mathcal{H}$ . This is an obvious consequence of (5); in addition,  $\omega \geq 1$ , because on  $V_0$  both forms coincide.

*Stable decomposition* We have to prove that there exist  $C_0$  independent of  $h, H, \mathcal{H}, \rho$ , and a decomposition of  $u \in V_h$ ,

$$u = \sum_{i=0}^{N_H} u^{(i)} \quad \text{with } u^{(i)} \in V_i, \quad (13)$$

such that

$$\sum_{i=0}^{N_H} A_h(u^{(i)}, u^{(i)}) \leq C_0^2 \mathcal{A}_h(u, u) \quad \forall u \in V_h. \quad (14)$$

In order to construct the stable decomposition of  $u \in V_h$ , we first define  $u^{(0)} \in V_0$  such that on each  $D_n \in \mathcal{T}_{\mathcal{H}}, n = 1, \dots, N_{\mathcal{H}}$ ,

$$u|_{D_n}^{(0)} = \bar{u}_{(n)} = \frac{1}{|D_n|} \int_{D_n} u \, dx.$$

Note that  $u^{(0)}$  is constant on the local solvers' subdomains  $\Omega_i, i = 1, \dots, N_H$ , as well.

Next, for  $i = 1, \dots, N_H$ , we define  $u^{(i)} \in V_i$ , as a zero-extension of the restriction of  $u - u^{(0)}$  to  $\Omega_i$ :

$$u^{(i)} = \begin{cases} u - u^{(0)} & \text{on } \Omega_i, \\ 0 & \text{elsewhere.} \end{cases}$$

We obviously have, since  $u^{(0)}$  is subdomainwise constant and its jumps cannot occur inside any element of  $\mathcal{T}_H$ ,

$$\begin{aligned} \sum_{i=1}^{N_H} A_h(u^{(i)}, u^{(i)}) &= \sum_{i=1}^{N_H} \sum_{K \in \mathcal{T}_h(\Omega_i)} (\rho \nabla u, \nabla u)_K + \sum_{i=1}^{N_H} \sum_{e \in \mathcal{E}_h^0(\Omega_i)} \gamma |[u]|_{0,e}^2 \\ &\quad + \sum_{i=1}^{N_H} \sum_{e \in \mathcal{E}_h^\partial(\Omega_i)} \gamma |(u - u^{(0)})_i|_{0,e}^2 \\ &\leq A_h(u, u) + \sum_{i=1}^{N_H} \sum_{e \in \mathcal{E}_h^\partial(\Omega_i)} \gamma |(u - u^{(0)})_i|_{0,e}^2. \end{aligned}$$

Let us consider subdomain  $\Omega_i$ . Then, it follows that on  $e \in \mathcal{E}_h^\partial(\Omega_i)$  such that  $e = \partial K_I \cap \partial K_J$  and  $K_I \in \mathcal{T}_h(\Omega_i), K_J \in \mathcal{T}_h(\Omega_j)$  there holds

$$\gamma = \frac{\delta}{h_e} \frac{\rho|_{K_I} \rho|_{K_J}}{\rho|_{K_I} + \rho|_{K_J}} \lesssim \frac{\delta}{h_{K_I}} \rho|_{K_I} \lesssim \frac{r_i}{h_i}$$

and analogously, if  $e$  lies on the boundary of  $\Omega$ , then

$$\gamma = \frac{\delta \rho|_{K_I}}{h_e} \lesssim \frac{r_i}{\underline{h}_i},$$

so that

$$\sum_{e \in \mathcal{E}_h^\partial(\Omega_i)} \gamma |(u - u^{(0)})_i|_{0,e}^2 \lesssim \frac{r_i}{\underline{h}_i} \sum_{e \in \mathcal{E}_h^\partial(\Omega_i)} |(u - u^{(0)})_i|_{0,e}^2 = \frac{r_i}{\underline{h}_i} |(u - u^{(0)})_i|_{0,\partial\Omega_i}^2.$$

From (A2) it follows that we can apply the trace inequality of Lemma 3 to  $u - u_i^{(0)}$  and obtain

$$\begin{aligned} & |(u - u^{(0)})_i|_{0,\partial\Omega_i}^2 \\ & \lesssim \frac{1}{H_i} |u - u^{(0)}|_{\Omega_i}^2 + H_i \left( \sum_{K \in \mathcal{T}_h(\Omega_i)} |\nabla u|_{0,K}^2 + \sum_{e \in \mathcal{E}_h^0(\Omega_i)} \frac{1}{h_e} |[u]|_{0,e}^2 \right), \end{aligned}$$

because again  $[u_i^{(0)}]$  vanishes on the inner skeleton  $\mathcal{E}_h^0(\Omega_i)$ . Then,

$$\sum_{i=1}^{N_H} \sum_{e \in \mathcal{E}_h^\partial(\Omega_i)} \gamma |(u - u^{(0)})_i|_{0,e}^2 \lesssim \sum_{i=1}^{N_H} \frac{r_i}{\underline{h}_i H_i} |u - u^{(0)}|_{\Omega_i}^2 + \max_{i=1,\dots,N_H} \frac{r_i}{\underline{r}_i} \frac{H_i}{\underline{h}_i} A_h(u, u).$$

Now,

$$\begin{aligned} \sum_{i=1}^{N_H} \frac{r_i}{\underline{h}_i H_i} |u - u^{(0)}|_{\Omega_i}^2 &= \sum_{n=1}^{N_{\mathcal{H}}} \sum_{i: \Omega_i \subset D_n} \frac{r_i}{\underline{h}_i H_i} |u - u^{(0)}|_{\Omega_i}^2 \\ &\leq \sum_{n=1}^{N_{\mathcal{H}}} \frac{R_n}{\underline{h} H_n} |u - \bar{u}_{(n)}|_{D_n}^2. \end{aligned}$$

Applying Poincaré's inequality for discontinuous finite element functions [4, Theorem 5.1] (see also [8, 14]) on the matching submesh of  $D_n$  and then using the shape- and contact-regularity assumption followed by the scaling argument we get

$$\begin{aligned} |u - \bar{u}_{(n)}|_{D_n}^2 &\lesssim \mathcal{H}_n^2 \left( \sum_{K \in \mathcal{T}_h(D_n)} |\nabla u|_{0,K}^2 + \sum_{e \in \mathcal{E}_h^0(D_n)} \frac{1}{h_e} |[u]|_{0,e}^2 \right) \\ &\lesssim \mathcal{H}_n^2 \left( \sum_{i: \Omega_i \subset D_n} \sum_{K \in \mathcal{T}_h(\Omega_i)} |\nabla u|_{0,K}^2 + \sum_{e \in \mathcal{E}_h^0(D_n)} \frac{1}{h_e} |[u]|_{0,e}^2 \right). \end{aligned}$$

Because on  $e \in \mathcal{E}_h^0(D_n)$  such that  $e = \partial K_I \cap \partial K_J$  there holds  $\min\{\rho_{K_I}, \rho_{K_J}\} \leq 2\rho_{K_I}\rho_{K_J}/(\rho_{K_I} + \rho_{K_J})$ , we conclude that

$$\underline{R}_n \leq \frac{2\gamma h_e}{\delta} \quad \text{on } \mathcal{E}_h^0(D_n),$$

and therefore

$$\begin{aligned} |u - \bar{u}_{(n)}|_{D_n}^2 &\lesssim \mathcal{H}_n^2 \left( \sum_{i:\Omega_i \subset D_n} \frac{1}{\underline{r}_i} \sum_{K \in \mathcal{T}_h(\Omega_i)} \rho_{|K|} |\nabla u|_{0,K}^2 + \frac{1}{\underline{R}_n} \sum_{e \in \mathcal{E}_h^0(D_n)} \gamma |[u]|_{0,e}^2 \right) \\ &\lesssim \mathcal{H}_n^2 \frac{1}{\underline{R}_n} \sum_{i:\Omega_i \subset D_n} \left( \sum_{K \in \mathcal{T}_h(\Omega_i)} \rho_{|K|} |\nabla u|_{0,K}^2 + \sum_{e \in \mathcal{E}_h(\Omega_i)} \gamma |[u]|_{0,e}^2 \right). \end{aligned}$$

From the above we can estimate

$$\sum_{i=1}^{N_H} \sum_{e \in \mathcal{E}_h^\partial(\Omega_i)} \gamma |(u - u^{(0)})_i|_{0,e}^2 \lesssim \left( \max_{n=1,\dots,N_H} \frac{\mathcal{H}_n^2}{hH_n} \frac{R_n}{\underline{R}_n} + \max_{i=1,\dots,N_H} \frac{H_i}{\underline{h}_i} \frac{r_i}{\underline{r}_i} \right) A_h(u, u)$$

and in consequence

$$\sum_{i=1}^{N_H} A_h(u^{(i)}, u^{(i)}) \lesssim \beta A_h(u, u) \lesssim \beta \mathcal{A}_h(u, u),$$

the latter inequality following from Lemma 1.

It remains to bound  $A_h(u^{(0)}, u^{(0)})$ . We have

$$A_h(u^{(0)}, u^{(0)}) = \sum_{e \in \mathcal{E}_h} \gamma |[u^{(0)}]|_{0,e}^2 \lesssim \sum_{e \in \mathcal{E}_h} \gamma |[u - u^{(0)}]|_{0,e}^2 + \sum_{e \in \mathcal{E}_h} \gamma |[u]|_{0,e}^2$$

For any  $e \in \mathcal{E}_h$  such that  $e = \partial K_I \cap \partial K_J$  and  $K_I \in \mathcal{T}_h(\Omega_i)$  and  $K_J \in \mathcal{T}_h(\Omega_j)$  we have

$$|[u - u^{(0)}]|_{0,e}^2 \lesssim |(u - u^{(0)})_i|_{0,e}^2 + |(u - u^{(0)})_j|_{0,e}^2,$$

so that

$$\sum_{e \in \mathcal{E}_h} \gamma |[u^{(0)}]|_{0,e}^2 \lesssim \sum_{i=1}^{N_H} \sum_{e \in \mathcal{E}_h^\partial(\Omega_i)} \gamma |(u - u^{(0)})_i|_{0,e}^2 + A_h(u, u)$$

Since the first term has already been estimated by  $\beta A_h(u, u)$ , we conclude that also

$$A_h(u^{(0)}, u^{(0)}) \lesssim \beta \mathcal{A}_h(u, u),$$

which completes the proof.

**Remark 1** In (A1) or (A2) we could have assumed the existence of a finite set of several such reference structures, the number of which is another absolute constant independent of  $h$ ,  $H$ ,  $\mathcal{H}$  and the jumps in  $\rho$ .

Below we prove that the convergence speed of the method is independent of the magnitude of jumps of  $\rho$ , if the variation of coefficient  $\rho$  is assumed moderate:

- either on the skeleton of the local solvers' partition  $\mathcal{E}_H$ ,
- or on the coarse partition  $\mathcal{T}_\mathcal{H}$ .

Observe that then large jumps of  $\rho$  are allowed, respectively, either on islands strictly inside local subdomains or on the boundaries of the coarse partition.

**Corollary 1** *In addition to the assumptions of Theorem 1 suppose that the variation of the coefficient  $\rho$  on the skeleton  $\mathcal{E}_H$  is uniformly bounded by an absolute constant of order 1:*

$$\rho|_e \simeq 1 \quad \forall e \in \mathcal{E}_H. \quad (15)$$

Then

$$\text{cond}(T) = O \left( \max_{n=1, \dots, N_\mathcal{H}} \frac{\mathcal{H}_n^2}{hH_n} + \max_{i=1, \dots, N_H} \frac{H_i}{h_i} \right).$$

*Proof* Since  $\rho$  is normalized so that  $\rho \geq 1$  it follows that  $1 \leq \underline{r}_i$  and  $1 \leq \underline{R}_n$ . Therefore  $1/\underline{R}_n \leq 1$  and, since  $\mathcal{E}_h^\partial(\Omega_i) \subset \mathcal{E}_H$  for  $i = 1, \dots, N_H$ ,

$$\frac{r_i}{\underline{r}_i} \leq r_i \lesssim 1,$$

so in consequence  $\max_n R_n \lesssim 1$ . □

**Corollary 2** *In addition to the assumptions of Theorem 1 suppose that the variation of the coefficient  $\rho$  is uniformly bounded by an absolute constant within elements of the coarse partition  $\mathcal{T}_\mathcal{H}$ :*

$$\rho|_{D_n} \simeq \bar{\rho}_{(n)} := \frac{1}{|D_n|} \int_{D_n} \rho \, dx, \quad n = 1, \dots, N_\mathcal{H}, \quad (16)$$

(for example,  $\rho$  can be piecewise constant on  $\mathcal{T}_\mathcal{H}$  with arbitrary jumps across the interface  $\mathcal{E}_\mathcal{H}$ ). Then

$$\text{cond}(T) = O \left( \max_{n=1, \dots, N_\mathcal{H}} \frac{\mathcal{H}_n^2}{hH_n} + \max_{i=1, \dots, N_H} \frac{H_i}{h_i} \right).$$

*Proof* By our assumption, for any  $i = 1, \dots, N_H$  and for  $n$  such that  $\Omega_i \subset D_n$  we have

$$r_i, \underline{r}_i \simeq \bar{\rho}_{(n)},$$

so that  $r_i/\underline{r}_i \lesssim 1$ , and for any  $n = 1, \dots, N_{\mathcal{H}}$  there holds

$$R_n, \underline{R}_n, \simeq \bar{\rho}_{(n)}$$

as well.  $\square$

**Remark 2** If we suppose that (15) or (16) is satisfied and that all meshes  $\mathcal{T}_h$ ,  $\mathcal{T}_H$  and  $\mathcal{T}_{\mathcal{H}}$  are quasi-uniform, then (with a little abuse of notation) the expression for  $\text{cond}(T)$  simplifies to

$$\text{cond}(T) = O\left(\frac{\mathcal{H}^2}{hH}\right),$$

where  $h$ ,  $H$ ,  $\mathcal{H}$  here mean the *maximum* diameter of elements in  $\mathcal{T}_h$ ,  $\mathcal{T}_H$  and  $\mathcal{T}_{\mathcal{H}}$ , respectively.

If each local solvers' subdomain reduces to a single element of the fine mesh, i.e.  $\mathcal{T}_H = \mathcal{T}_h$ , one has to solve in parallel  $N_h = O(h^d)$  local problems, which are just a very simple  $3 \times 3$  (in 2-D) or  $4 \times 4$  (in 3-D) system each, while the coarse problem remains fixed with (relatively small) size  $N_{\mathcal{H}} = O(\mathcal{H}^d)$ . The price to be paid for this massive parallelism is slower (as compared to the standard approach, cf. Corollary 4 below) convergence of the iteration:

**Corollary 3** *In addition to our general assumptions of Sect. 2, assume that  $\mathcal{T}_H = \mathcal{T}_h$  and (A1) holds together with (16). Then*

$$\text{cond}(T) = O\left(\max_{n=1, \dots, N_{\mathcal{H}}} \frac{\mathcal{H}_n^2}{\min_{K \in \mathcal{T}_h(D_n)} h_K^2}\right).$$

*Proof* Assumption (A2) is automatically satisfied since  $\Omega_i$  reduces to single element in  $\mathcal{T}_h$ , so the estimate follows directly from Corollary 2.  $\square$

Theorem 1 also recovers previously known bounds, for constant or piecewise constant coefficient, in the “standard” case when  $\mathcal{T}_H = \mathcal{T}_{\mathcal{H}}$  (see [1, 10, 14]), and extends them to the case of coefficient mildly varying over  $\mathcal{E}_H$ :

**Corollary 4** *In addition to our general assumptions of Sect. 2, assume that  $\mathcal{T}_H = \mathcal{T}_{\mathcal{H}}$  and (A2) holds together with either (15) or (16). Then*

$$\text{cond}(T) = O\left(\max_{i=1, \dots, N_H} \frac{H_i}{\min_{K \in \mathcal{T}_h(\Omega_i)} h_K}\right).$$

*Proof* The proof follows straightforwardly from Corollaries 1 and 2 and thus is omitted.  $\square$

For completeness, let us also mention the following:

**Corollary 5** *In addition to our general assumptions of Sect. 2, assume that  $\mathcal{T}_h = \mathcal{T}_H = \mathcal{T}_{\mathcal{H}}$ . Then  $\text{cond}(T) = O(1)$ .*

*Proof* Assumption (16) is automatically fulfilled on a single element, while (A1) and (A2) are trivially satisfied because of our assumptions on the fine grid in Sect. 2. The estimate then follows again from Corollary 2.  $\square$

Although the condition number in Corollary 5 is independent both from  $h$  and the jumps of  $\rho$ , the preconditioned operator  $T$  is not robust in this particular case: the number of degrees of freedom in the coarse space operator  $T_0$  is unacceptably large, of order  $N_h$ , i.e. of the same order as the original size of the discrete problem. On the other hand,  $T_0$  is defined only on the space of piecewise constant functions.

## 4 Numerical experiments

Let us choose the unit square  $[0, 1]^2$  as the domain  $\Omega$  and for some prescribed integer  $\mathcal{M}$  divide it into  $N_{\mathcal{H}} = 2^{\mathcal{M}} \times 2^{\mathcal{M}}$  smaller squares  $D_n$  ( $n = 1, \dots, N_{\mathcal{H}}$ ) of equal size. This coarse grid  $\mathcal{T}_{\mathcal{H}}$  is then refined into a uniform local solvers'  $2^M \times 2^M$  square grid  $\mathcal{T}_H$ , which in turn is further refined into a uniform fine triangulation  $\mathcal{T}_h$  based on a square  $2^m \times 2^m$  grid ( $m \geq M \geq \mathcal{M}$ ) with each square split into two triangles of identical shape. Hence, the grid parameters are  $h = 2^{-m}$ ,  $H = 2^{-M}$  and  $\mathcal{H} = 2^{-\mathcal{M}}$ ; obviously, there holds  $\mathcal{T}_{\mathcal{H}} \subseteq \mathcal{T}_H \subseteq \mathcal{T}_h$ . We discretize problem (1) on the fine mesh  $\mathcal{T}_h$  using (4) with  $\delta = 7$ .

In the following tables we report the number of Preconditioned Conjugate Gradient iterations for operator  $T$  which are required to reduce the initial Euclidean norm of the residual by a factor of  $10^6$  and (in parentheses) the condition number of  $T$  approximation computed from the extreme eigenvalues estimate based on the PCG coefficients, see e.g. [15, Sect. 6.7.3]. We always choose a random vector for the right hand side and zero as the initial guess.

### 4.1 Dependence on $h$ , $H$ , $\mathcal{H}$

First, let us consider the convergence of the “massively parallel” method when  $\mathcal{T}_H = \mathcal{T}_h$  against “standard” approach, when  $\mathcal{T}_H = \mathcal{T}_{\mathcal{H}}$ . For the diffusion coefficient we take an elementwise  $P_0$  approximation of a continuous function  $\rho(x) = x_1^2 + x_2^2 + 1$ . As it turns out from Table 1, the condition number of the method considered in Corollary 3

**Table 1** Dependence of the number of iterations and the condition number (in parentheses) on  $\mathcal{H} = 2^{-\mathcal{M}}$  and  $h = 2^{-m}$  for the method with  $\mathcal{T}_H = \mathcal{T}_h$

Coarse ( $\mathcal{M}$ )	Fine ( $m$ )			
	4	5	6	7
4	29 (22)	39 (40)	59 ( $1.1 \times 10^2$ )	96 ( $3.8 \times 10^2$ )
5		30 (23)	39 (40)	59 ( $1.1 \times 10^2$ )
6			30 (23)	38 (40)
7				30 (23)

Slowly varying coefficient  $\rho$



**Table 2** Dependence of the number of iterations and the condition number (in parentheses) on  $\mathcal{H} = 2^{-\mathcal{M}}$  and  $h = 2^{-m}$  for the method with  $\mathcal{T}_H = \mathcal{T}_{\mathcal{H}}$ 

	Coarse ( $\mathcal{M}$ )	Fine ( $m$ )			
		4	5	6	7
Slowly varying coefficient $\rho$	4	27 (20)	35 (34)	46 (67)	62 ( $1.3 \times 10^2$ )
	5		28 (20)	35 (34)	46 (67)
	6			28 (20)	35 (34)
	7				28 (20)

**Table 3** Dependence of the number of iterations and the condition number (in parentheses) on the discontinuity when the coefficient is constant inside coarse partition and  $\mathcal{T}_H = \mathcal{T}_h$ 

$\rho_B$	$10^0$	$10^2$	$10^4$	$10^6$
Iter (cond)	134 ( $3.8 \times 10^2$ )	143 ( $3.7 \times 10^2$ )	160 ( $3.7 \times 10^2$ )	181 ( $3.7 \times 10^2$ )

Red–black  $4 \times 4$  distribution of  $\rho$ , aligned with the coarse grid. Fixed  $\mathcal{H}/h = 8$

indeed shows an  $O(\mathcal{H}^2/h^2)$  behavior, as predicted, while methods which use local solves on subdomains of diameter at least  $\mathcal{H}$  (e.g. [1, 14]) exhibit  $O(\mathcal{H}/h)$  dependence, cf. Corollary 4 and Table 2.

## 4.2 Strongly discontinuous coefficient piecewise constant on $\mathcal{T}_{\mathcal{H}}$

Here we deal with large discontinuities in  $\rho$  under restrictions of Corollary 3. The case of discontinuous coefficient when  $\mathcal{T}_H = \mathcal{T}_{\mathcal{H}}$  (cf. Corollary 4) has already been treated elsewhere [12], so we restrict ourselves only to the case when  $\mathcal{T}_H = \mathcal{T}_h$ . Let us then consider  $\rho$  with discontinuities aligned with an auxiliary partitioning of  $\Omega$  into  $4 \times 4$  squares. Precisely, we introduce a red–black checkerboard coloring of this partitioning and set  $\rho$  equal to  $\rho_R = 1$  in red regions, and the value of  $\rho_B$  reported in Table 3 in black ones. In this way, our fine and coarse triangulations, with  $m = 7$  and  $\mathcal{M} = 4$ , are always aligned with the discontinuities, as required in Corollary 3. Table 3 confirms the condition number is independent from  $\rho_B$  in this case.

## 4.3 Strongly discontinuous coefficient piecewise constant on $\mathcal{T}_h$

Next, we investigate elementwise discontinuous coefficient with  $m = 7$  and  $\mathcal{M} = 4$  as before, but with  $\rho = 1$  on odd and  $\rho = \rho_B$  on even-numbered triangles of  $\mathcal{T}_h$ . Table 4 shows that in this case the preconditioner fails (a dash means the method did not converge in 600 iterations) for large jumps in  $\rho$ . This confirms the importance of controlled variation of the coefficient inside elements over the partition of the domain.

As stated in Corollary 5, in the unfeasible case when  $\mathcal{T}_h = \mathcal{T}_H = \mathcal{T}_{\mathcal{H}}$ , the condition number of the preconditioned system is  $O(1)$  regardless of  $h$  and the jumps of  $\rho$  (parameters  $H$  and  $\mathcal{H}$  do not apply here) and this is confirmed in experiments, as reported in Table 5.

**Table 4** Dependence of the number of iterations and the condition number (in parentheses) on the discontinuity when the coefficient is elementwise discontinuous

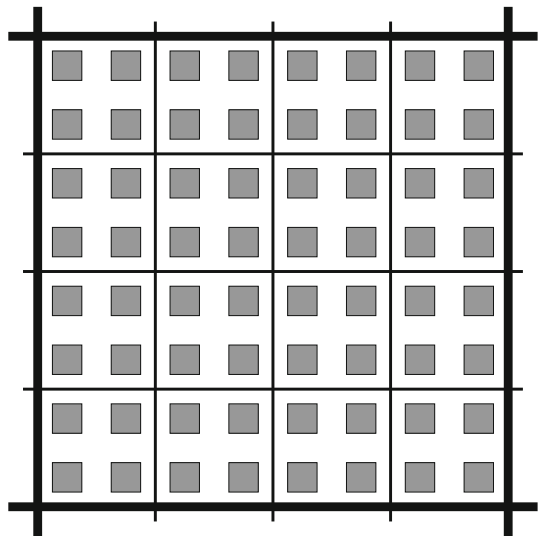
$\rho_B$	$10^0$	$10^2$	$10^4$	$10^6$
Iter (cond)	134 ( $3.8 \times 10^2$ )	435 ( $3.8 \times 10^3$ )	— ( $3.1 \times 10^5$ )	— ( $2.5 \times 10^7$ )

Fixed  $\mathcal{H}/h = 8$  and  $\mathcal{T}_H = \mathcal{T}_h$

**Table 5** Dependence of the number of iterations and the condition number (in parentheses) on the discontinuity and fine mesh size  $h = 2^{-m}$  when the coefficient is elementwise discontinuous and  $\mathcal{T}_h = \mathcal{T}_H = \mathcal{T}_{\mathcal{H}}$ 

$\rho_B$	Fine ( $m$ )				
	3	4	5	6	7
$10^0$	40 (34)	44 (41)	45 (42)	47 (44)	47 (45)
$10^6$	38 (22)	41 (24)	42 (25)	66 (50)	64 (50)

**Fig. 2** Example distribution of coefficient  $\rho$ , considered in Sect. 4.4, depicted inside a single square interior element of  $\mathcal{T}_{\mathcal{H}}$  (whose edges are marked with *thick lines*) when  $\mathcal{M} = 2$ . In gray islands  $\rho = 10^6$ ; elsewhere  $\rho = 1$ . Local solvers' grid marked with *thinner lines* for  $M = 4$ . Note that in presented example,  $\rho$  on  $\mathcal{E}_H$  is constant



#### 4.4 Strongly discontinuous coefficient constant on $\mathcal{E}_H$

In Corollary 1 we show that in some circumstances one can allow very large variation of  $\rho$  even inside local solvers' subdomains. Let us then choose  $\rho$  constant equal to 1, except for square islands, where we set  $\rho = 10^6$ . The islands are centrally located in cells of a  $2^5 \times 2^5$  independent division of  $\Omega$  into squares. The size of every island equals half-length of the cell (see Fig. 2 for an example). In order to make the fine triangulation resolve the discontinuities of  $\rho$  well, we restrict ourselves only to  $m = 8$ .

As it turns out from Table 6, if local solvers' subdomains entirely cover the islands—that is, if  $M \leq 5$ —then  $\rho$  turns constant on  $\mathcal{E}_H$  and the performance of

**Table 6** Dependence of the number of iterations and the condition number (in parentheses) on  $\mathcal{H} = 2^{-\mathcal{M}}$  and  $H = 2^{-M}$  when the coefficient is discontinuous with contrast ratio  $10^6$  on small islands centrally placed on a  $2^5 \times 2^5$  grid, cf. Fig. 2

Coarse ( $\mathcal{M}$ )	Local ( $M$ )							
	2	3	4	5	6	7	8	
2	132 ( $4.9 \cdot 10^2$ )	151 ( $5.8 \cdot 10^2$ )	188 ( $8.6 \cdot 10^2$ )	253 ( $1.4 \cdot 10^3$ )	$-(5.6 \cdot 10^8)$	$-(5.1 \cdot 10^8)$	$-(1.4 \cdot 10^9)$	
3		120 ( $3.1 \cdot 10^2$ )	129 ( $3.5 \cdot 10^2$ )	152 ( $4.7 \cdot 10^2$ )	$-(1.6 \cdot 10^8)$	$-(1.4 \cdot 10^8)$	$-(3.7 \cdot 10^8)$	
4			93 ( $1.7 \cdot 10^2$ )	98 ( $1.8 \cdot 10^2$ )	737 ( $4.7 \cdot 10^7$ )	751 ( $4.3 \cdot 10^7$ )	883 ( $1.1 \cdot 10^8$ )	
5				69 ( $9.0 \cdot 10^1$ )	76 ( $1.1 \cdot 10^2$ )	81 ( $1.1 \cdot 10^2$ )	88 ( $1.3 \cdot 10^2$ )	
6					77 ( $1.2 \cdot 10^2$ )	99 ( $1.3 \cdot 10^2$ )	115 ( $1.5 \cdot 10^2$ )	
7						62 ( $5.2 \cdot 10^1$ )	59 ( $4.6 \cdot 10^1$ )	
8								

Fine grid parameter is  $h = 2^{-m}$  with  $m = 8$

the preconditioner is not sensitive to large jumps in the coefficients inside subdomains, which is in agreement with Corollary 1. Moreover, in the case when there are jumps across  $\mathcal{E}_H$  and at the same time the coarse grid does not resolve the discontinuities well enough, i.e.  $M > 5 > \mathcal{M}$ , the condition number becomes sensitive to the coefficient jumps.

When  $\mathcal{M} \geq 5$ , we are back to the case when the coefficient is constant on  $\mathcal{T}_H$  (i.e. the case of Corollary 2) and results from Table 6 reflect this.

## 5 Conclusions

A nonoverlapping ASM preconditioner for symmetric interior penalty DG discretization of second order elliptic PDE with discontinuous coefficient on a nonmatching mesh  $\mathcal{T}_h$  has been presented. A very large number of small local problems defined by subdomains introduced by a domain partition  $\mathcal{T}_H$  is solved in parallel, together with one coarse problem of moderate size on a coarser partition  $\mathcal{T}_{\mathcal{H}}$  such that  $\mathcal{T}_{\mathcal{H}} \subseteq \mathcal{T}_H \subseteq \mathcal{T}_h$  in the sense of subsequent refinements. Two cases of arrangement of the coefficient discontinuity relative to partitionings  $\mathcal{T}_{\mathcal{H}}$  or  $\mathcal{T}_H$  have been identified, for which the condition number of the resulting system is  $O(\mathcal{H}^2/hH)$  independently of the jumps of the coefficient. This property allows one to design a massively parallel method with rate of convergence independent of the jumps of the coefficient. Our result also shows that the decrease of local solvers' subdomains size  $H$  (i.e. the increase of parallelism) affects the condition number only linearly, while it turns out highly important to keep the diameter  $\mathcal{H}$  of the coarse grid as small as possible, due to its quadratic influence on the condition number.

**Acknowledgments** We would like to thank an anonymous referee whose comments and remarks helped to improve the paper. This research has been partially supported by the Polish National Science Centre Grant 2011/01/B/ST1/01179.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

1. Antonietti, P.F., Ayuso, B.: Schwarz domain decomposition preconditioners for discontinuous Galerkin approximations of elliptic problems: non-overlapping case. *M2AN Math. Model. Numer. Anal.* **41**(1), 21–54 (2007)
2. Ayuso de Dios, B., Holst, M., Zhu, Y., Zikatanov, L.: Multilevel preconditioners for discontinuous Galerkin approximations of elliptic problems with jump coefficients. *Math. Comp.* **83**, 1083–1120 (2014)
3. Bjørstad, P.E., Dryja, M., Vainikko, E.: Additive Schwarz methods without subdomain overlap and with new coarse spaces. In: *Domain Decomposition Methods in Science and Engineering* (Beijing, 1995), pp. 141–157. Wiley, Chichester (1997)
4. Brenner, S.C.: Poincaré-Friedrichs inequalities for piecewise  $H^1$  functions. *SIAM J. Numer. Anal.* **41**(1), 306–324 (2003)
5. Brix, K., Pinto, C., Canuto, C., Dahmen, W.: Multilevel preconditioning of discontinuous-Galerkin spectral element methods part I: Geometrically conforming meshes. *IGPM Preprint*, RWTH Aachen (2013)

6. Canuto, C., Pavarino, L.F., Pieri, A.B.: BDDC preconditioners for continuous and discontinuous Galerkin methods using spectral/ $hp$  elements with variable local polynomial degree. *IMA J. Numer. Anal.* **34**(3), 879–903 (2014)
7. Crouzeix, M., Raviart, P.-A.: Conforming and nonconforming finite element methods for solving the stationary Stokes equations. I. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge* **7**(R-3), 33–75 (1973)
8. Di Pietro, D.A., Ern, A.: Discrete functional analysis tools for discontinuous Galerkin methods with application to the incompressible Navier-Stokes equations. *Math. Comp.* **79**(271), 1303–1330 (2010)
9. Di Pietro, D.A., Ern, A.: Mathematical aspects of discontinuous Galerkin methods. In: *Mathématiques & Applications (Berlin) [Mathematics & Applications]*, vol. 69. Springer, Heidelberg (2012)
10. Dryja, M.: On discontinuous Galerkin methods for elliptic problems with discontinuous coefficients. *Comput. Methods Appl. Math.* **3**(1):76–85 (electronic), (2003)
11. Dryja, M., Krzyżanowski, P.: Additive Schwarz methods for DG discretization of elliptic problems with discontinuous coefficient. In: *Domain Decomposition Methods in Science and Engineering XXII. Lecture Notes in Computer Science and Engineering*. Springer, Berlin (to appear, 2015)
12. Dryja, M., Krzyżanowski, P., Sarkis, M.: Additive Schwarz method for DG discretization of anisotropic elliptic problems. In: Erhel, J., Gander, M.J., Halpern, L., Pichot, G., Sassi, T., Widlund, O. (eds.) *Domain Decomposition Methods in Science and Engineering XXI. Lecture Notes in Computational Science and Engineering*, vol. 98, pp. 407–415. Springer, Berlin (2014)
13. Ern, A., Stephansen, A.F., Zunino, P.: A discontinuous Galerkin method with weighted averages for advection-diffusion equations with locally small and anisotropic diffusivity. *IMA J. Numer. Anal.* **29**(2), 235–256 (2009)
14. Feng, X., Karakashian, O.A.: Two-level additive Schwarz methods for a discontinuous Galerkin approximation of second order elliptic problems. *SIAM J. Numer. Anal.*, **39**(4):1343–1365 (electronic) (2001)
15. Saad, Y.: Iterative methods for sparse linear systems. In: *Society for Industrial and Applied Mathematics*, 2nd edn., Philadelphia, PA (2003)
16. Toselli, A., Widlund, O.: Domain decomposition methods—algorithms and theory. In: *Springer Series in Computational Mathematics*, vol. 34. Springer, Berlin (2005)